AD_____

Award Number: DAMD17-00-1-0031

TITLE: Functional Analysis of C-CAM1 Tumor Suppressor Gene by
Targeted Gene Deletion

PRINCIPAL INVESTIGATOR: Sue-Hwa Lin, Ph.D.

CONTRACTING ORGANIZATION: The University of Texas M.D. Anderson
Cancer Center
Houston, Texas 77030

REPORT DATE: January 2004

TYPE OF REPORT: Final Addendum

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are
those of the author(s) and should not be construed as an official
Department of the Army position, policy or decision unless so
designated by other documentation.

BEST AVAILABLE COPY

20040830 057

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 074-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE January 2004	3. REPORT TYPE AND DATES COVERED Final Addendum (1 Jan 2003 – 31 Dec 2003)	
4. TITLE AND SUBTITLE Functional Analysis of C-CAM1 Tumor Suppressor Gene by Targeted Gene Deletion			**5. FUNDING NUMBERS** DAMD17-00-1-0031
6. AUTHOR(S) Sue-Hwa Lin, Ph.D.			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) The University of Texas M.D. Anderson Cancer Center Houston, Texas 77030 **E-Mail:** slin@notes.mdacc.tmc.edu			**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012			**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**

11. SUPPLEMENTARY NOTES
Original contains color plates: All DTIC reproductions will be in black and white.

12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for Public Release; Distribution Unlimited	12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 Words)

CEACAM1 is a cell adhesion molecule. Alternative splicing of the Ceacam1 gene generated two isoforms of CEACAM1, i.e. CEACAM1-L (long form) and CEACAM1-S (short form). We have shown that CEACAM1-L plays critical roles in prostate cancer progression. We propose to determine the function of CEACAM1 in prostate development, prostate homeostasis and prostate tumorigenesis by using gene targeting technologies. Specifically, we plan to generate mice with a targeted deletion of the CEACAM1 cytoplasmic domain. We have constructed a Ceacam1 targeting vector in which exons 7-9 of Ceacam1 gene was flanked by loxP sites for the generation of conditional knockout of the Ceacam1 gene. Embryonic stem cell clones containing the recombinant gene allele were established and injected into blastocysts for germ line transmission of the targeting construct. We have succeeded in producing mice harboring the conditional Ceacam1 knockout allele. In this extended period, we generated mice with homozygote conditional Ceacam1 allele and characterized the expression of CEACAM1 protein. In addition, we have crossed the homozygote conditional Ceacam1 mice with transgenic mice containing probasin-driven cre to generate mice with prostate-specific deletion of Ceacam1 gene. Initial analysis of 3 month old mouse prostates showed that deletion of Ceacam1-L in the mouse prostates led to epithelial hyperplasia.

14. SUBJECT TERMS Prostate cancer			15. NUMBER OF PAGES 23
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT Unlimited

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

Table of Contents

(4)   INTRODUCTION

C-CAM1 (renamed as CEACAM1) is a cell adhesion molecule of the immunoglobulin supergene family (1). We have shown that CEACAM1 plays critical roles in prostate cancer initiation and progression and that loss of CEACAM1 is an early event in the development of prostate cancer (2). Although tumorigenesis studies in mouse xenograft model have suggested the involvement of CEACAM1 in epithelial cell growth and differentiation, the functional roles of CEACAM1 in normal prostate development, prostate homeostasis, and prostate tumorigenesis remain unclear. Towards the aim of unraveling the roles of CEACAM1's growth suppressive activity in prostate growth and tumorigenesis, we propose to use gene targeting and embryonic stem cell technologies to generate CEACAM1 knockout mice. Specifically, we plan (1) to determine the roles of CEACAM1's growth suppressive function in vivo by generating mice with a targeted deletion of the CEACAM1 cytoplasmic domain; (2) to determine the roles of CEACAM1's growth suppressive function in prostate development and tumorigenesis by generating mice with a prostate-specific knockout of the CEACAM1 cytoplasmic domain. The proposed work was divided into two Tasks to be carried out in parallel.

Task 1. Generate mice with targeted deletion of CEACAM1 cytoplasmic domain to determine the roles of CEACAM1's growth suppressive function in vivo (months 1-30)

Task 2. Prostate-specific loss of function of CEACAM1 gene in prostate (months 7-36)

Genetic manipulation of mouse genes in vivo is a powerful approach to understanding the function of a gene, both during embryonic development and in adult tissues. This method requires full knowledge of the genomic structure of the gene of interest. Unlike humans and rats, which each have one CEACAM1 gene, two CEACAM1-like genes, *Ceacam1* and *Ceacam2*, were identified in mice. In the previous study, we have isolated and sequenced these two closely related *Ceacam* genes from a mouse 129 Sv/Ev library (3). We have also examined the tissue-specific and embryonic expressions of these mouse *Ceacam1* and *Ceacam2* genes (3). Our sequence analysis revealed that the genes encoded nine exons and spanned approx. 16-17 kb (*Ceacam1*) and 25 kb (*Ceacam2*). The genes were highly similar (79.6%). The major differences in the protein-coding regions were located in exons 2, 5 and 6. To determine whether

4

functional redundancy exists between *Ceacam1* and *Ceacam2*, we examined their expression in 16 mouse tissues by using semi-quantitative reverse transcription-PCR. As in human and rat, in mouse *Ceacam1* mRNA was highly abundant in the liver, small intestine, prostate and spleen. In contrast, *Ceacam2* mRNA was only detected in kidney, testis and, to a lesser extent, spleen. In a mouse embryo, *Ceacam1* mRNA was detected at day 8.5, disappeared between days 9.5 and 12.5, and re-appeared at day 19. On the other hand, no *Ceacam2* mRNA was detected throughout embryonic development. The different tissue expression patterns and regulation during embryonic development suggest that the *Ceacam1* and *Ceacam2* proteins, although highly similar, may have different functions both during mouse development and in adulthood. Results from this study allow us to design a gene targeting strategy that is specific to *Ceacam1* gene and also allow us to perform both straight and conditioned knockout of *Ceacam1* gene in parallel.

We have designed a gene targeting strategy that is specific to *Ceacam1* gene. In addition, the *Ceacam1* gene in the targeting vector was flanked by loxP sites to allow for generation of both straight and conditional knockout of *Ceacam1* gene. The targeting vector has been constructed and 24 embryonic stem cell clones containing the recombinant gene allele have been established. Two of the embryonic cell clones were injected into blastocysts for germ line transmission of the targeting construct toward the end of second funding period. In the final year of funding period, we have succeeded in producing mice harboring conditional *Ceacam1* gene. Two founder mice from two embryonic cell clones showed germ line transmission of the targeting construct. Mice with homozygote conditional *Ceacam1* allele were born from these two founder mice suggesting that embryonic lethality does not occur in mice carrying conditional *Ceacam1* allele. We requested a one-year no cost extension to continue the project. In this extended one year, we generated sufficient mice with homozygote conditional Ceacam1 allele and characterized the expression of CEACAM1 protein in the wild-type and homozygote conditional Ceacam1 mice. In addition, we have further crossed the homozygote conditional Ceacam1 mice with transgenic mice containing probasin-driven cre (PB-Cre4) to generate mice with prostate-specific deletion of Ceacam1 gene. Initial analysis of 3-month-old mouse prostates showed that deletion of Ceacam1 in the mouse prostates led to epithelial hyperplasia. These results suggest that CEACAM1 has a critical role in prostate growth.

(5) BODY (Progress report)

Because Task 1 and Task 2 are performed in parallel, we describe the progresses in these aspects together.

5. 1. Gene targeting strategy

Based on genomic characterization of *Ceacam1* gene, we decided to delete the cytoplasmic domain of CEACAM1. This strategy is based on the following rationales: (1) the entire *Ceacam1* gene, around 20 kb, is too large to delete; and (2) the cytoplasmic domain of CEACAM1 is critical for tumor suppression. A conditional knockout construct is designed to delete exons 7~9 in a tissue-specific or non-specific manner in the mice. ES cells harboring the targeting vector are injected into blastocysts to generate mice carrying the conditional *Ceacam1* allele. For straight knockout, the goal of Task 1, the conditional knockout mice will be crossed with transgenic mice carrying actin-driven Cre recombinase, which is expressed during embryonic development. For conditional knockout, the goal of Task 2, the *Ceacam1* gene will be deleted in prostate specifically by crossing the conditional knockout mice with transgenic mice carrying probasin-driven Cre recombinase. Thus, Task 1 and Task 2 are being performed in parallel. The gene knockout strategy is shown in Fig. 1.

Fig. 1. Maps of the *Ceacam1* genomic locus and the loxP-targeting vector designed to introduce a Neo^R cassette. Exons are denoted by rectangles and loxP sites are denoted by triangles. Red rectangles indicate CEACAM coding regions. Successful targeting yields a conditional Ceacam1^{floxed} allele. Cre-mediated recombination results in deletion of Neo^R cassette and the cytoplasmic domain exons, yielding a knockout allele. Specific primers used to screen for various alleles are indicated by small arrows labeled a through c.
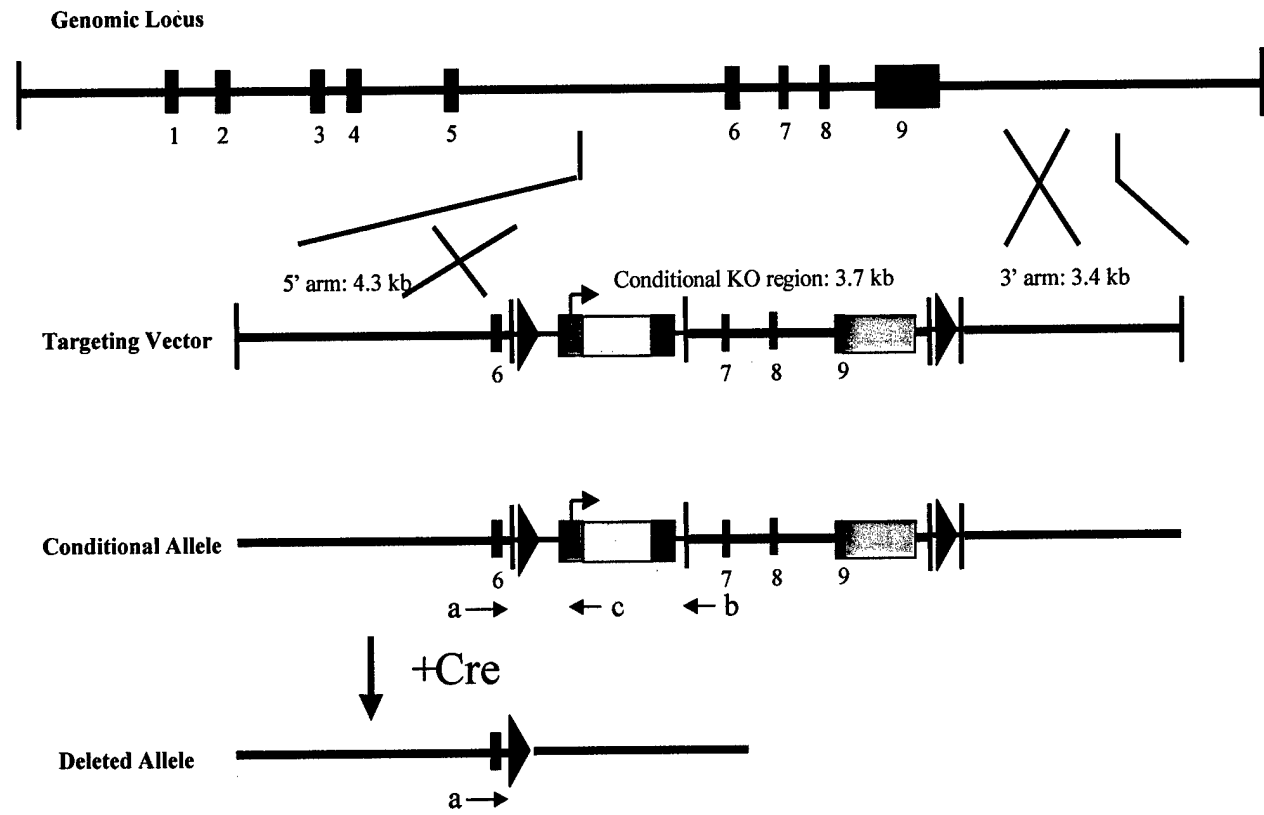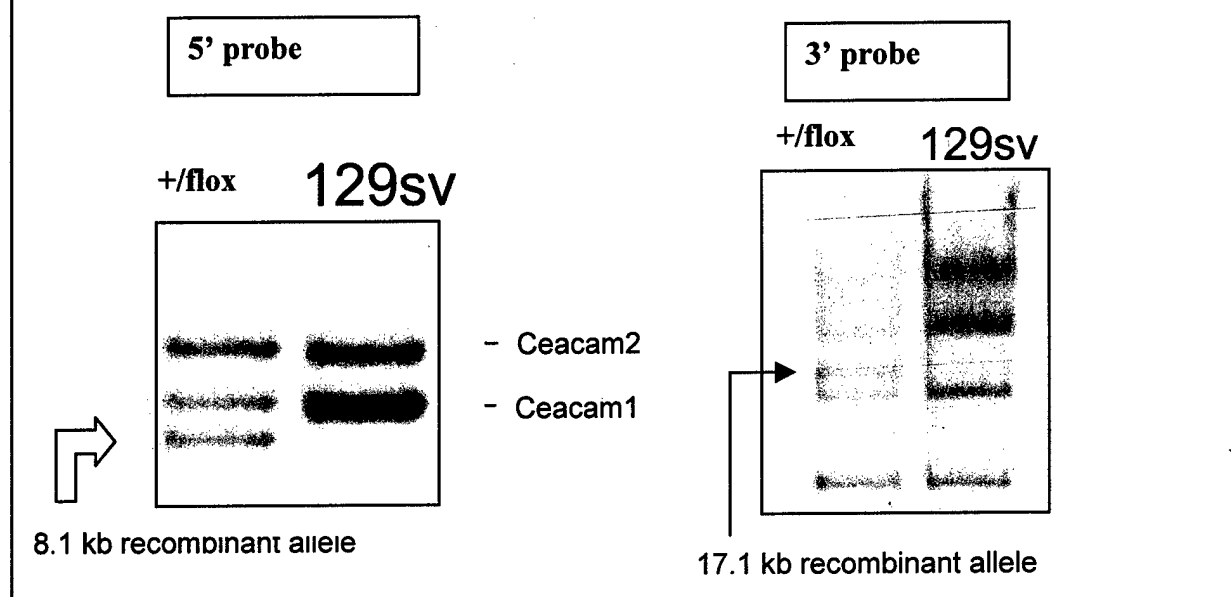
5. 2. Targeting vector construction

A 129Sv/Ev mouse genomic library in BAC vectors was screened with a probe generated from cDNA coding for the full-length mouse CEACAM1. Three positive BAC clones were identified and their structures were determined by restriction mapping. Consistent with our previous study (3), two closely related genes, i.e. *Ceacam1* and *Ceacam2*, were identified from restriction mapping. Based on tissue-specific and embryonic expression of *Ceacam1* and *Ceacam2* gene (3), we decided to delete the cytoplasmic domain of *Ceacam1*. A conditional knockout construct was designed to delete exons 7~9. The conditional targeting construct was designed by inserting a loxP-neo cassette at the Xmn1 site in intron 6 and a loxP fragment at the HindIII site in intron 9 of *Ceacam1* (Fig. 1). The targeting vector includes 4.3 kb of homologous DNA upstream of the loxP-neo site and 3.4 kb of homologous DNA downstream of the second loxP site. The targeting vector was verified by restriction mapping.

5. 3. Transfection and generation of embryonic stem cells harboring targeting vector

The targeting vector was transfected into embryonic stem cells by electroporation. A total of 270 ES cell clones were obtained from two electroporations. Genomic DNA was extracted from these ES cell clones. Half of the DNA from each sample was digested with SphI and processed for Southern blot analysis. The 5' probe used is a 0.7kb ApaI/SpeI fragment from subclone BamHI-A. Using the 5' probe, a ~9.2 kb band corresponding to the endogenous *Ceacam1* and a ~10 kb band corresponding to *Ceacam2* were detected in most samples. Many clones (a total of 24) included an additional hybridization signal at the size (8.1 kb) predicted for the homologous recombinant allele (Fig. 2). These potential positives clones were digested with Xho I and hybridized with the 3' probe, which is a 1.2kb Sac I/Xho I fragment from subclone Apa I-4. A ~15.3 kb band corresponding to the endogenous allele of Ceacam1 and a ~17.1 kb band corresponding to recombinant allele were detected in 16 clones (Fig. 2). These 16 clones were also detected with a Neo probe, which gave a 17.1 kb band in Southern blot analysis. Taken together, a total of 16 ES cell clones containing the *Ceacam1* conditional allele were generated. Three positive ES cell clones were injected into mouse blastocysts using procedures described in Chang et al. (4). Chimeric mice were obtained and they were mated with C57BL mice. The F1 mice having germ line transmission of the loxP-targeted *Ceacam1* allele were interbred to generate F2 mice.

Fig. 2. Southern blot analysis performed with a 5' and a 3' external probes to detect homologous recombination that gives rise to conditional Ceacam1flox allele.

5' probe

3' probe

+/flox 129sv

+/flox 129sv

- Ceacam2
- Ceacam1

8.1 kb recombinant allele

17.1 kb recombinant allele

5. 4. Genotyping of the wild-type and loxP-targeted (floxed) alleles

Germ line transmission was confirmed by PCR analysis. Isolated mouse tail DNA was used for PCR identification of the genotype. Three sets of primers that detect the wild-type and the *floxed* allele of Ceacam1, and the Ceacam2 gene were used to genotype the mice. Primer A (5'   ACACAAGGAGGCCTCTCAGATGGCG   3')   and   Primer   C (5'GCGCCTCCCCTACCCGGTAGAATT 3'), containing sequences from Exon 6 of *Ceacam1* and neomycin cassette, respectively, will produce a 488 bp PCR product from the *floxed Ceacam1* allele. Primer A and Primer B (5' GACTTTGGCTTCCTGACTGGAGGA 3'), containing sequences from Intron 6, will generate a 382 bp PCR product from wild-type *Ceacam1*. Two founder mice from two embryonic cell clones showed germ line transmission of the targeting construct containing the conditional *Ceacam1* gene. Mice with homozygous conditional *Ceacam1* alleles were born from these two founder mice. These results suggest that embryonic lethality does not occur in the mice carrying conditional *Ceacam1* knockout alleles, thus it is possible to pursue the function of CEACAM1 in normal prostate development, prostate homeostasis and prostate tumorigenesis *in vivo*.

9

Fig. 3. PCR analysis with mouse tail DNA to detect homologous recombination in mice.

Flox/flox   +/flox   +/+

488 b.p. →

382 b.p. →
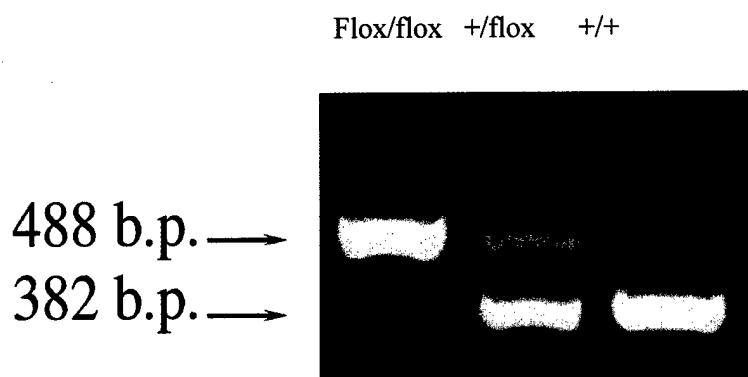
5. 5. CEACAM1 protein expression in wild-type and CEACAM1^{flox/flox} mouse

To examine the effect of flox allele on the expression of CEACAM1 protein, we performed Western blot analysis on the tissue lysates prepared from prostates of wild-type and CEACAM1$^{flox/flox}$ mice. We found that the floxed CEACAM1 allele led to decreased expression of CEACAM1 protein (Fig. 4). Such a hypomorphic phenotype is commonly seen in loxP-targeted allele. Previous studies have established that there are two alternatively spliced form of CEACAM1, i.e. CEACAM1-L (long form) and CEACAM1-S (short form). Because only CEACAM1-L has tumor suppressive activity, our gene targeting strategy was to delete the cytoplasmic domain of CEACAM1-L. We examined the effect of the conditional Ceacam1 allele on the expression of CEACAM1-L and CEACAM1-S. Due to the glycosylation of CEACAM1, the presence of CEACAM1-L and CEACAM1-S, which have molecular mass difference of 7 kDa, cannot be resolved well in SDS-PAGE directly. We thus assessed the amount of CEACAM1-L and CEACAM1-S in the wild-type and CEACAM1$^{flox/flox}$ mice by treating the tissue lysates with PNGase to remove N-linked glycosylation. As shown in Fig. 4, CEACAM1-L and CEACAM1-S are present in both wild-type and CEACAM1$^{flox/flox}$ cell lysates. However, only CEACAM1-L showed significant decrease in CEACAM1$^{flox/flox}$ mice when compared with those in wild-type mice.

Fig. 4. CEACAM1 flox/flox mouse is hypomorphic. A and B represent two set of experiments with different wild-type and Ceacam1$^{flox/flox}$ mice.
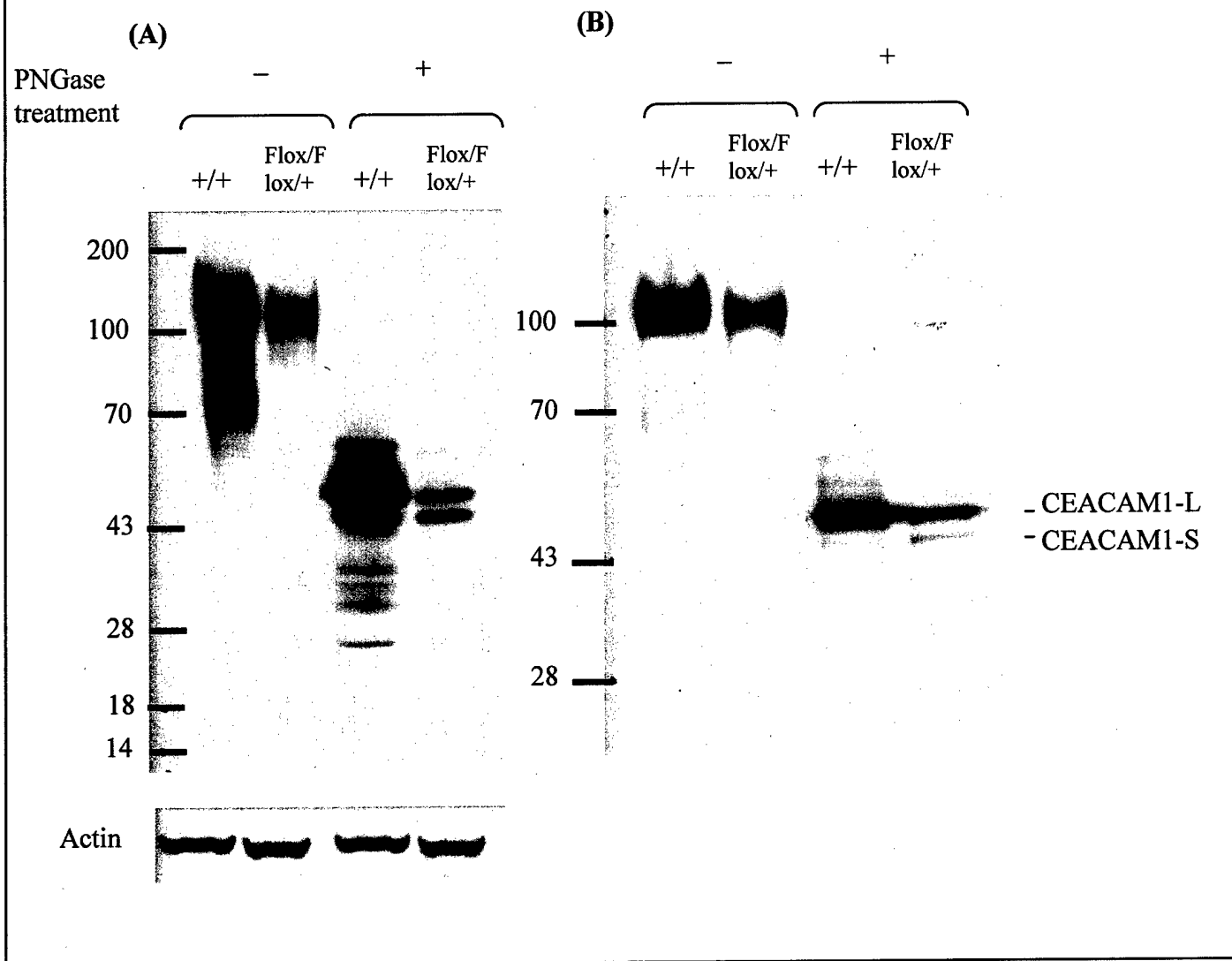
5, 6. Effect of cre-recombinase on Ceacam1 conditional allele.

We test whether the loxp sites that flank exons 7, 8, and 9 of Ceacam1 allele can be removed by cre recombinase by using recombinant adenovirus with cre gene (Ad-Cre). CEACAM1$^{flox/flox}$ mice were treated with Ad-Cre or control adenovirus Ad-gal (containing beta-galactosidase gene) by i.v. injection of adenovirus. Western blots were performed on the cell lysates prepared from livers of these mice at 21 days post-injection. As shown in Fig. 5, there is a significant decrease in the level of CEACAM1-L in the mouse treated with Ad-Cre. This

observation suggests that exons 7-9 of Ceacam1 can be removed by treatment of CEACAM1$^{flox/flox}$ mice with cre recombinase.

Fig. 5. Removal the CEACAM1-L cytoplasmic domain by Cre recombinase
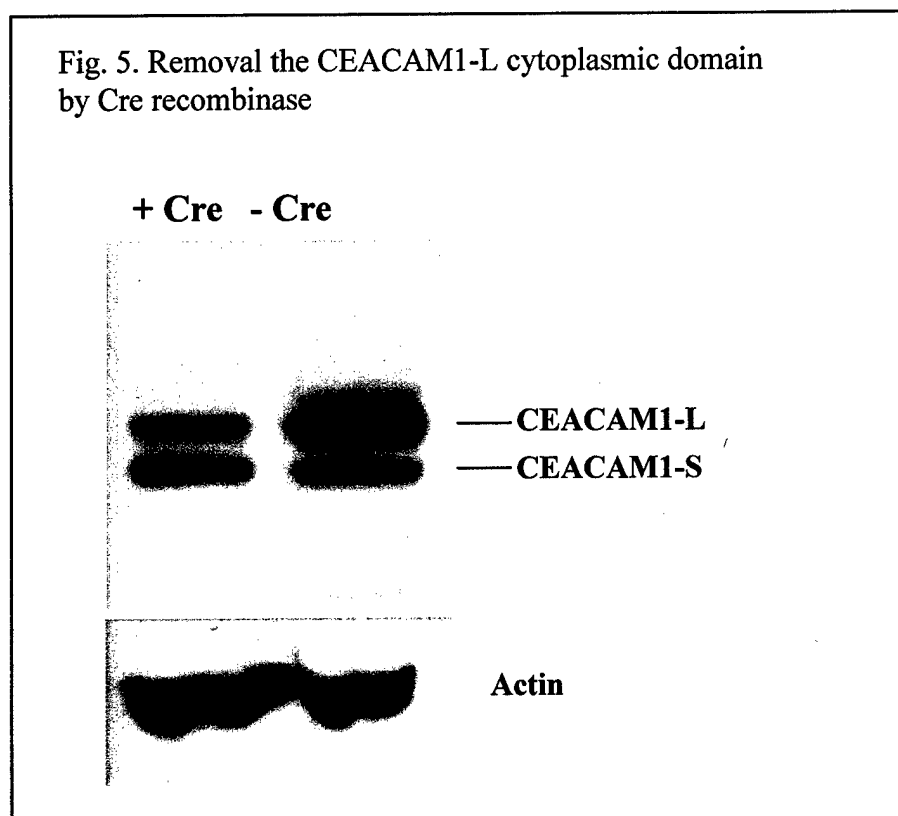
5. 7. Generation of prostate-specific deletion of Ceacam1 by crossing CEACAM1$^{flox/flox}$ mice with PB-Cre4 mice

To generate mice with prostate-specific deletion of Ceacam1, we crossed the CEACAM1$^{flox/flox}$ mice with the transgenic mouse (PB-cre4) from NIH mouse consortium. PB-cre4 mice is a transgenic mouse line containing cre under the ARR2PB promoter. ARR2PB is a third generation probasin promoter that has been designed to target high levels of transgene expression specifically to the prostate. This new composite PB promoter construct (*ARR2PB*) is less than 500 bp in size, but is androgen and glucocorticoid regulated and prostate-specific, and it can target high level expression of a protein in the prostate of the transgenic mice (5). CEACAM1$^{flox/flox}$ mice containing cre allele was confirmed by PCR analysis. Isolated mouse-tail DNA was used for PCR using primer Cre-F (5' ttgcctgcattaccggtcgatgca 3') and Cre-R (5'

gatcctggcaatttcggctat 3') and the PCR product is 500 bp. Male mice homozygote with Ceacam1 floxed allele with (Ceacam1$^{flox/flox}$;cre+) and without (Ceacam1$^{flox/flox}$;cre-) cre were selected for phenotype analysis. We have obtained two mice each with the specific genotype. Because probasin promoter becomes active at 5-7 weeks after birth, we killed these mice at the age of three months and the prostates were collected and analyzed. As shown in Fig. 6, prostates from mice without cre transgene showed relatively normal morphology. In the mice with cre transgene, there was a decrease in the size of the luminal epithelial cells and these cells seem to be less differentiated compared to the prostates from mice without cre. The morphological changes were observed in both ventral prostate and dorsal prostate. We are expanding these mouse colonies to characterize the effect of loss of CEACAM1-L on prostate.

Fig. 6. Morphology of dorsal prostates.

Ceacam1$^{flox/flox}$;cre-   Ceacam1$^{flox/flox}$;cre+
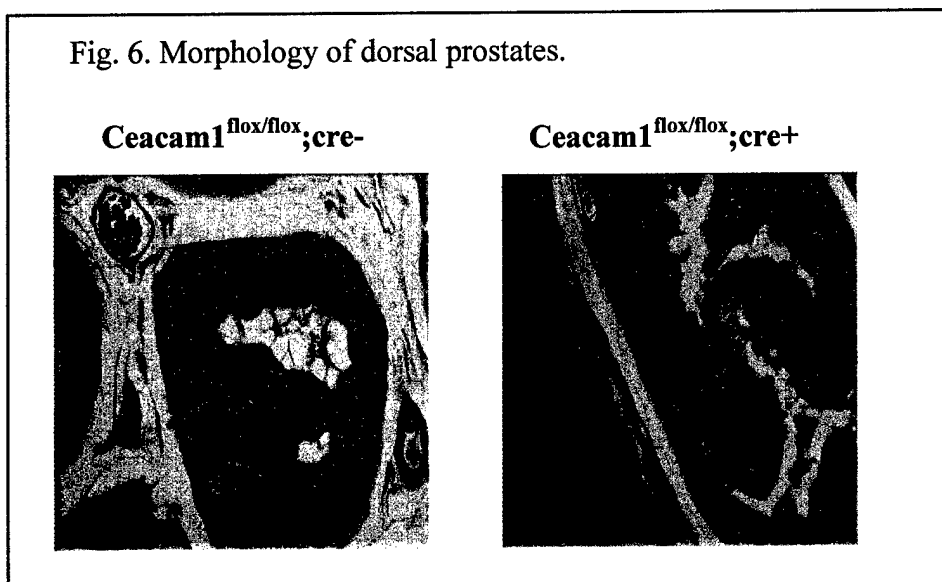
(6) KEY RESEARCH ACCOMPLISHMENTS

❖ Designed a knockout strategy that allows Task1 and Task2 to be carried out in parallel.
❖ Constructed the targeting vector.
❖ Transfected the targeting vector into embryonic stem cells and selected 24 positive ES cell clones.
❖ Injected three ES cell clones into blastocysts.

- ❖ Bred and genotyped mice for the presence of recombinant allele.
- ❖ Obtained mice carrying heterozygote and homozygote recombinant allele.
- ❖ Analyzed the expression of CEACAM1 protein in mice carrying floxed Ceacam1 allele to test for the effect of floxed allele on CEACAM1 protein expression
- ❖ Analyzed the effect of cre recombinase on the floxed exons.
- ❖ Cross the mice with homozygote Ceacam1 floxed allele with transgenic mice carrying probasin promoter driven cre gene.
- ❖ Analyzed the prostate phenotype of mice homozygote with Ceacam1 floxed allele with and without probasin-driven cre transgene.

(7) REPORTABLE OUTCOMES

Han, E., Phan, D., Lo, P., Poy, M. N., Behringer, R., Najjar, S., and Lin, S.-H.: Difference in tissue-specific and embryonic expression of mouse Ceacam1 and Ceacam2 genes. Biochem. J. (2001) 355, 417-423.

(8) CONCLUSION
We propose to determine the roles of CEACAM1's growth suppressive activity in prostate growth and tumorigenesis by using gene targeting and embryonic stem cell technologies to generate knockout mice. We have designed and constructed a gene targeting vector for specific disruption of the *Ceacam1* gene. This gene targeting strategy will allow us to perform both straight and conditioned knockout of *Ceacam1* gene in parallel. Embryonic stem cells containing the knockout construct were generated and injected into blastocysts. We have succeeded in generating mice carrying *Ceacam1* conditional knock out allele. We found that the recombinant allele alters the endogenous *Ceacam1* expression level. Deletion of the C-terminal portion of Ceacam1 gene in the mouse prostate by probasin promoter driven cre gene resulted in less differentiated luminal epithelial cells in the mouse prostate at 3 month of age.

(9) REFERENCES

1. Beauchemin, N. and Lin, S.-H. Role of C-CAM as a tumor suppressor. Amsterdam: Harwood Academic Publishers, 1998.

2. Lin, S.-H. and Pu, Y.-S. Function and therapeutic implication of C-CAM cell-adhesion

molecule in prostate cancer. Seminars in Oncology, *26:* 227-233, 1999.

3. Han, E., Phan, D., Lo, P., Poy, M. N., Behringer, R., Najjar, S. M., and Lin, S.-H. Differences in tissue-specific and embryonic expression of mouse Ceacam1 and Ceacam2 genes. Biochem. J., *355:* 417-423, 2001.

4. Chang, A. S. Y., Reardon, M. J., and DeMayo, F. J. Germline genetic engineering techniques in endocrinology. Totowa, NJ: Humana Press Inc., 2000.

5. Zhang, J., Gao, N., Kasper, S., Reid, K., Nelson, C., and Matusik, R. J. An androgen-dependent upstream enhancer is essential for high levels of probasin gene expression. Endocrinology, *145:* 134-148, 2004.

List of personnel receiving/received pay from the research effort:
Sue-Hwa Lin
Karen Earley
Weiping Luo

Differences in tissue-specific and embryonic expression of mouse *Ceacam1* and *Ceacam2* genes

Eric HAN*, Dillon PHAN*, Piao LO*, Matthew N. POY†, Richard BEHRINGER‡, Sonia M. NAJJAR† and Sue-Hwa LIN*[1]

*Department of Molecular Pathology, The University of Texas M. D. Anderson Cancer Center, Houston, TX 77030, U.S.A., †Department of Pharmacology and Therapeutics, Medical College of Ohio, Toledo, OH 43614, U.S.A., and ‡Department of Molecular Genetics, The University of Texas M. D. Anderson Cancer Center, Houston, TX 77030, U.S.A.

The intercellular adhesion molecule CEACAM1, also known as C-CAM1 (where CAM is cell-adhesion molecule), can function as a tumour suppressor in several carcinomas, including those of the prostate, breast, bladder and colon. This suggests that CEACAM1 may play an important role in the regulation of normal cell growth and differentiation. However, there is no direct evidence to support this putative function of CEACAM1. To elucidate its physiological function by targeted gene deletion, we isolated the *Ceacam* genes from a mouse 129 Sv/Ev library. Although there is only one *Ceacam1* gene in humans and one in rats, two homologous genes (*Ceacam1* and *Ceacam2*) have been identified in the mouse. Our sequence analysis revealed that the genes encoded nine exons and spanned approx. 16–17 kb (*Ceacam1*) and 25 kb (*Ceacam2*). The genes were highly similar (79.6%). The major differences in the protein-coding regions were located in exons 2, 5 and 6 (76.9%, 87.0% and 78.5% similarity respectively). In addition, introns 2, 5 and 7 were also significantly different, being 29.7%, 59.8% and 64.5% similar respectively. While most of these differences were due to nucleotide substitutions, two insertions of 418 and 5849 bp occurred in intron 2 of *Ceacam2*, and another two insertions of 1384 and 197 bp occurred in introns 5 and 7 respectively. To determine whether functional redundancy exists between *Ceacam1* and *Ceacam2*, we examined their expression in 16 mouse tissues by using semi-quantitative reverse transcription–PCR. As in human and rat, in the mouse *Ceacam1* mRNA was highly abundant in the liver, small intestine, prostate and spleen. In contrast, *Ceacam2* mRNA was only detected in kidney, testis and, to a lesser extent, spleen. Reverse transcription–PCR using testis RNA indicated that *Ceacam2* in the testis is an alternatively spliced form containing only exons 1, 2, 5, 6, 8 and 9. In the mouse embryo, *Ceacam1* mRNA was detected at day 8.5, disappeared between days 9.5 and 12.5, and re-appeared at day 19. On the other hand, no *Ceacam2* mRNA was detected throughout embryonic development. The different tissue expression patterns and regulation during embryonic development suggest that the CEACAM1 and CEACAM2 proteins, although highly similar, may have different functions both during mouse development and in adulthood.

Key words: C-CAM, cell-adhesion molecule, gene expression, tumour suppressor.

INTRODUCTION

The intercellular adhesion molecule CEACAM1, also known as C-CAM1 [1] (where CAM is cell-adhesion molecule), is a member of the Ig gene family [2]. In addition to its cell-adhesion function, CEACAM1 is also a tumour suppressor in prostate [3], breast [4], bladder [5] and colon [6] carcinomas. Consistent with its role as a tumour suppressor, loss of CEACAM1 expression was observed in hepatomas [7], colon carcinomas [8–10] and endometrial [11] and prostate [12,13] cancers. This suggests that CEACAM1 may regulate cell growth and differentiation.

To elucidate the physiological functions of CEACAM1, we plan to generate mice with targeted *Ceacam* deletions. Genetic manipulation of mouse genes *in vivo* is a powerful approach for understanding the function of a gene, both during embryonic development and in adult tissues. This method requires full knowledge of the genomic structure of the gene of interest. Unlike humans and rats, which each have one *Ceacam1* gene [14–16], two *Ceacam1*-like genes, *Ceacam1* and *Ceacam2*, were identified in BALB/c mice [17]. This poses potential problems in any attempt to manipulate these genes. First, if the two genes have the same function, deletion of one gene may not produce any phenotypic alteration. Secondly, if there is no significant difference in the sequences of the genes, any genetic manipulation may not be specific for a single gene, or it might be difficult to ascertain which gene was altered. It is, therefore, essential to determine the complete sequences and expression profiles of these related *Ceacam* genes in the 129 Sv mouse strain before genetic manipulation is performed.

We report herein the isolation and sequencing of the full-length mouse *Ceacam1* and *Ceacam2* genes. They both encode nine exons, with significant sequence differences in certain regions. These sequence differences will allow selective targeting of one *Ceacam* gene compared with another. In addition, the tissue-specific distributions of CEACAM1 and CEACAM2 were different, and only *Ceacam1* was expressed during embryonic development. The differences in their tissue and developmental expression patterns suggest that these two genes may have different functions *in vivo*.

MATERIALS AND METHODS

Isolation and characterization of *Ceacam1* and *Ceacam2*

A 129 Sv/Ev mouse genomic library in λLEX (provided by Dr Li-Yuan Yu-Lee, Baylor College of Medicine, Houston, TX,

---

U.S.A.) was screened with the 1.6 kb full-length mouse *Ceacam1* cDNA [18], which had been labelled by using the Klenow fragment of DNA polymerase I, random hexanucleotide primers and [α-^{32}P]dCTP [19]. A total of 24 positive clones were identified in the initial screening of this genomic library. PCR with Oligo 127 (5′-GTGTCACTCTAGGCTACAGGAAAT-3′) and Oligo 122 (5′-GAGGCCAGCTCCATGTCTCTGCTG-3′), which are specific to the 5′ region of mouse *Ceacam1* [20], showed that seven of the clones had the N-terminal *Ceacam1* sequence. Similarly, PCR with Oligo 120 (5′-GAAGTCTGGCGGATCT-GGCTCCTT-3′) and Oligo 131 (5′-TTGAAGTTCAGGACA-GTGTATGCG-3′) showed that seven clones had the 3′ *Ceacam1* sequence. These clones were isolated by secondary and tertiary screening and characterized by restriction mapping. The nucleotide sequences of the exons and introns were determined by primer walking using specific oligonucleotide primers. Sequencing was performed by the DNA Sequencing Core Facilities at M. D. Anderson Cancer Center with an automated fluorescent DNA sequencer (Applied Biosystems Inc., Ramsey, NJ, U.S.A.).

Analysis of *Ceacam1* and *Ceacam2* expression using a mouse cDNA panel

A mouse cDNA panel containing first-strand cDNA prepared from mouse tissues and normalized for β-actin expression was purchased from OriGene Technologies Inc. (Rockville, MD, U.S.A.) and used to analyse the expression of *Ceacam1* and *Ceacam2* in various mouse tissues. Oligonucleotides specific to exon 2 of *Ceacam1* (Oligo 116, 5′-AATCTGCCCCTGGCG-CTTGGAGCC-3′; Oligo 179, 5′-AAATCGCACAGTCGCC-TGAGTACG-3′) and to exon 2 of *Ceacam2* (Oligo 117, 5′-AATATGATGAAGGGAGTCTTGGCC-3′; Oligo 180, 5′-AAATTGTCCAGTCAGGACCCTACG-3′) were used as primers to detect specific mRNAs for *Ceacam1* and *Ceacam2* by PCR. PCR cycling conditions were as follows: (1) pre-denaturation (94 °C, 3 min) for one cycle; (2) denaturation (94 °C, 30 s), annealing (60 °C, 30 s) and extension (72 °C, 2 min) for 35 cycles; and (3) final extension (72 °C, 5 min). The predicted size of the PCR products was 246 bp. These PCR products were analysed by agarose-gel electrophoresis, transferred on to a nylon membrane and hybridized with ^{32}P-labelled oligonucleotide probes specific to *Ceacam1* (Oligo 181; 5′-AACAC-TACGGCTATAGACAAA-3′) and *Ceacam2* (Oligo 182; 5′-TCTACTACGTCTACAAATGCT-3′).

Ceacam2 cDNA sequence from mouse testis

RNA was prepared from mouse testis by using RNAzol B (TEL-TEST Inc. Friendswood, TX, U.S.A.). The cDNAs coding for *Ceacam2* were obtained from the testis RNA by reverse transcription–PCR (RT-PCR) with Oligo 630 (5′-GAATTCA-AGCTTAAGAAGCTAGCAGGCAGCAGAGAC-3′), which contains nucleotides −36 to +1 of exon 1, and Oligo 631 (5′-GCGGCCGCCTAATGATGATGATGATGATGATGCT-TCTTTTTTACTTCTGAATAAAC-3′), which is complementary to the end of the coding sequence in exon 9 plus seven histidine codons. RT-PCR was performed with Oligo 630 and Oligo 631 according to the procedures provided by the manufacturer (Amersham/Pharmacia Corp., Arlington Heights, IL, U.S.A.). The PCR product was subcloned into pCRII-topo (Invitrogen, San Diego, CA, U.S.A.) and its sequence was determined.

RESULTS

Isolation and characterization of mouse *Ceacam1* and *Ceacam2* genes

The seven clones with the 5′ sequence of *Ceacam1* had different restriction maps, as did another seven clones with the 3′ end. The different restriction maps seemed to reflect different genes. Grouping these clones according to their restriction maps and DNA sequences revealed two distinct sequences, representing *Ceacam1* and *Ceacam2*. The overlapping genomic clones that spanned *Ceacam1* and *Ceacam2* are shown in Figure 1. DNA sequence analysis of both strands of these two genes revealed that *Ceacam1* and *Ceacam2* each contain nine exons each, with sizes of approx. 18.3 and 24.65 kb respectively (Figure 1).

The sizes of the exons and introns and the intron/exon boundary sequences for *Ceacam1* and *Ceacam2* are shown in Tables 1 and 2 respectively. The first exon, which codes for the first two-thirds of the signal sequence, is 304 bp in both *Ceacam1* and *Ceacam2*. In contrast, exons 2–5 are each approx. 300 bp, and each codes for one Ig-like domain in both *Ceacam1* and *Ceacam2*. The transmembrane domain is encoded by exon 6 and the cytoplasmic domain by exons 7–9. Consistent with the RNA splicing rule [21], each intron starts with GT at the 5′ end and ends with AG at the 3′ end (Tables 1 and 2).

Comparison of *Ceacam1* and *Ceacam2* genes

The overall similarity between *Ceacam1* and *Ceacam2*, including all exons and introns, is approx. 79.6%. The similarities between
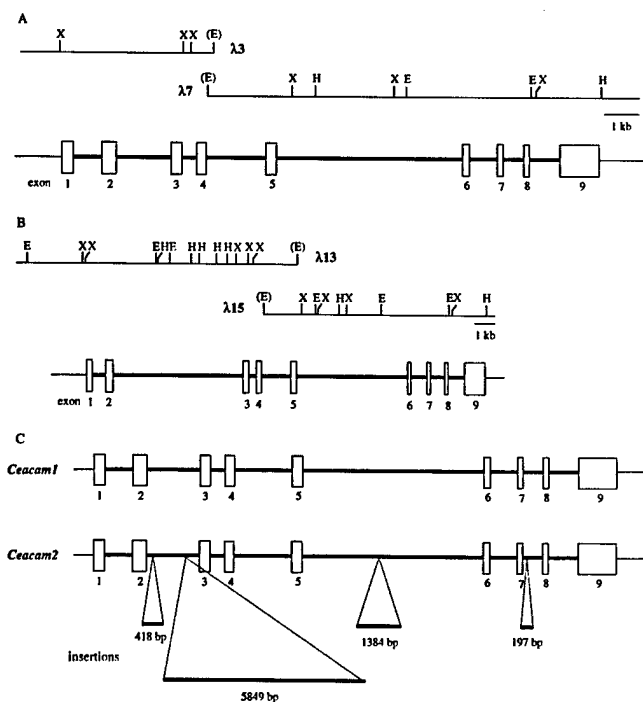
Figure 1   Structures of the mouse *Ceacam1* and *Ceacam2* genes

Numbered open boxes indicate exons. The thick lines connecting the exons represent introns. (**A**) The maps of two overlapping λ clones (λ3 and λ7) that contain *Ceacam1*. (**B**) The maps of two overlapping λ clones (λ13 and λ15) that contain *Ceacam2*. The restriction maps are aligned with the intron/exon map. E, *Eco*RI; H, *Hind*III; X, *Xba*I. The enzyme sites in parentheses are from the multiple cloning region of the λ phage vector. (**C**) Comparison of the mouse *Ceacam1* and *Ceacam2* genes.

Table 1   Intron/exon boundaries of *Ceacam1*

Exon	Size (bp)	Exon 3'			Intron	Size (bp)		Exon 5'	
1	304	ACA Thr	G A	gtaaggagatattcc ...	1	767	... tctcttccctcttag	CC la	TCA Ser
2	360	CAC His	C P	gtaagtaattatctg ...	2	1540	... ctattatctgcacag	CC ro	ATA Ile
3	285	ATC Ile	T T	gtgagtaacttcttc ...	3	467	... ttctgtttgctccag	AT yr	GGT Gly
4	255	CTT Leu	G G	gtaagtggatctctg ...	4	1727	... tctgtttgtccacag	AG ln	CCA Pro
5	276	ATA Ile	T P	gtgagtgacctcgcc ...	5	5358	... ttctttccctgacag	TT he	GAC Asp
6	121	GGC Gly	GG Gl	gtaggacagtctttc ...	6	796	... ctcatatttatttag	G y	GGA Gly
7	53	CAC His	A A	gtaagtaaagccaat ...	7	587	... ttcttcctcccctag	AT sn	CTG Leu
8	32	AAC Asn	AAG Lys	gtgagcactgccact ...	8	948	... ctctcatcctttcag	GTG Val	GAT Asp
9	1180								

Table 2   Intron/exon boundaries of *Ceacam2*

Exon	Size (bp)	Exon 3'			Intron	Size (bp)		Exon 5'	
1	304	ACA Thr	G A	gtaaggagatattcc ...	1	766	... tctcttccctcttag	CC la	TCA Ser
2	360	CAC His	A T	gtaagtaattctctg ...	2	7807	... acacagtgcagacag	CC hr	CTA Leu
3	285	ATC Ile	T T	gtgagtaacttcttt ...	3	465	... tcttgtttgctccag	AT yr	GGT Gly
4	255	CTT Leu	G G	gtaagtggatctctg ...	4	1717	... tctgtttgtccacag	AG lu	CCA Pro
5	276	ATA Ile	T P	gtgagtgaccttgcc ...	5	6737	... ttctttccctgacag	TT he	GAC Asp
6	118	CGC Arg	TG Tr	gtaggacagtctttc ...	6	798	... ctcatatttatttag	G p	GGA Gly
7	53	CAC His	A A	gtaagtaaagccaat ...	7	771	... ctcttcctcccctag	AT sn	CTG Leu
8	32	AAC Asn	AAG Lys	gtgagcactgccact ...	8	973	... ctctcatattttcag	GTG Val	GAT Asp
9	1180								

Table 3   Intron/exon lengths and similarities of *Ceacam1* and *Ceacam2*

	Length (bp)				Length (bp)		
Exon	*Ceacam1*	*Ceacam2*	Similarity (%)	Intron	*Ceacam1*	*Ceacam2*	Similarity (%)
1	304	304*	88.9	1	767	766	97.4
2	360	360	76.9	2	1540	7807	29.7
3	285	285	97.9	3	467	465	94.7
4	255	255	99.6	4	1726	1717	97.9
5	276	276	87.0	5	5358	6737	59.8
6	121	118	78.5	6	796	798	98.5
7	53	53	100	7	587	771	64.5
8	32	32	100	8	948	973	94.3
9	1180	1180†	86.8				

* Includes 240 bp of promoter and 5′ untranslated sequence.
† Includes 889 bp of 3′ untranslated sequence.

Exon 2 (nucleotide)

```
          843        853        863        873        883        893
Ceacam1  CCTCACTTTTAGCCTCCTGGAGCCCTGCCACCACTGCTGAAGTCACCATTGAGGCTGTGC
         ::::::::::::::::::::::::::::: :::::::::: :::: :: :: :::: : :
Ceacam2  CCTCACTTTTAGCCTCCTGGAGCCCTCCCACCACTGCACAAGTGACTGTTATGGCTTTTC
          840        850        860        870        880        890

          903        913        923        933        943        953
Ceacam1  CGCCCCAGGTTGCTGAAGACAACAATGTTCTTCTACTTGTTCACAATCTGCCCCTGGCGC
         : : ::: : :::::::::: ::::::::::: ::::: ::::: ::::: ::     ::
Ceacam2  CACTCCACGCTGCTGAAGGCAACAATGTTATTCTAGTTGTTTACAATATGATGAAGGGAG
          900        910        920        930        940        950

          963        973        983        993       1003       1013
Ceacam1  TTGGAGCCTTTGCCTGGTACAAGGGAA[ACACTACGGCTATAGACAAAG]AAATTGCACGAT
         :  :::::: :::: ::::::::: ::::::: ::: : :  ::::::  :::::
Ceacam2  TCTTGGCCTTTAGCTGGCACAAGGGAT[CTACTACGTCTACAAATGCTG]AAATTGTACGAT
          960        970        980        990       1000       1010

         1023       1033       1043       1053       1063       1073
Ceacam1  TTGTACCAAATAGTAATATGAATTTCACGGGGCAAGCATACAGCGGCAGAGAGATAATAT
         ::::: ::   : ::::: ::: : : :  : : :::: :::::::::: :::
Ceacam2  TTGTAACAGGCACTAATAAGACTATAAAAGGGCCTGTACACAGTGGCAGAGAGGCACTAT
         1020       1030       1040       1050       1060       1070

         1083       1093       1103       1113       1123       1133
Ceacam1  ACAGCAATGGATCCCTGCTCTTCCAAATGATCACCATGAAGGATATGGGAGTCTACACAC
         ::::::::::::::::::::: :::::: : :::::::::::::::::::::::::::::
Ceacam2  ACAGCAATGGATCCCTGCTCATCCAAAGGGTCACCATGAAGGATACGGGAGTCTACACAA
         1080       1090       1100       1110       1120       1130

         1143       1153       1163       1173       1183       1193
Ceacam1  TAGATATGACAGATGAAAACTATCGTCGTACTCAGGCGACTGTGCGATTTCATGTACACC
         :::: :::::::::: :: :::::::::::      ::::: : ::::::::::::::
Ceacam2  TAGAAATGACAGATCAAGACTATCGTCGTAGGGTCCTGACTGGACAATTTCATGTACACA
         1140       1150       1160       1170       1180       1190
```

Exon 2 (protein)

```
          32        42        52        62        72        82
Ceacam1  SLLASWSPATTAEVTIEAVPPQVAEDNNVLLLVHNLPLALGAFAWYKGNTTAIDKEIARF
         :::::::: ::: :: : :   :: ::::: : ::: :  :: : ::: :  :::
Ceacam2  SLLASWSPPTTAQVTVMAFPLHAAEGNNVILVVYNMMKGVLAFSWHKGSTTSTNAEIVRF
          32        42        52        62        72        82

          92       102       112       122       132
Ceacam1  VPNSNMNFTGQAYSGREIIYSNGSLLFQMITMKDMGVYTLDMTDENYRRTQATVRFHVH
         :  :   :      ::::: ::::::: : ::::: ::: ::: ::: : ::::
Ceacam2  VTGTNKTIKGPVHSGREALYSNGSLLIQRVTMKDTGVYTIXMTDQDYRRRVLTGQFHVH
          92       102       112       122       132
```
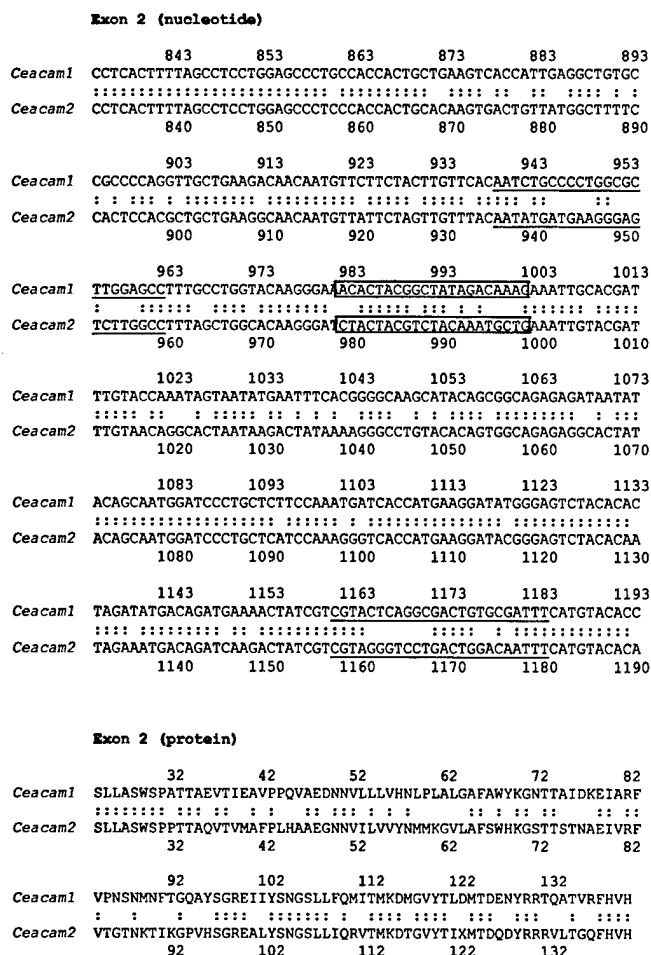
Figure 2 Comparison of the nucleotide and protein sequences of exon 2 of Ceacam1 and Ceacam2

Oligonucleotides for PCR analyses are underlined. Oligonucleotides for hybridization (Oligo 181 and Oligo 182) are boxed. The nucleotides are numbered with the A of the start ATG as the first nucleotide according to the genomic sequence. The amino acids are numbered with the start methionine (ATG) as the first amino acid.

Figure 3 Expression of Ceacam1 and Ceacam2 mRNAs in various mouse tissues

(A) Specificity of Oligos 181 and 182 in distinguishing Ceacam1 from Ceacam2. (B, C) Tissue-specific expression of Ceacam1 (B) and Ceacam2 (C) analysed by RT-PCR and hybridization with Oligo 181 and Oligo 182 respectively. The positions of the PCR products are indicated by the arrows on the right.

the exons and introns of *Ceacam1* and *Ceacam2* are shown in Table 3. The major differences between these two genes are in exon 2, intron 2, intron 5, exon 6 and intron 7, which are 76.9 %, 29.7 %, 59.8 %, 78.5 % and 64.5 % similar respectively. The nucleotide and the deduced amino acid sequences of exon 2 for *Ceacam1* and *Ceacam2* are shown in Figure 2. It is apparent that most of the sequence substitutions in exons 2, which code for the first Ig domain, do result in amino acid substitutions in *Ceacam1* and *Ceacam2*. Thus the amino acid identity between the first Ig domains of CEACAM1 and CEACAM2 is 57.1 %, which is significantly lower than the similarity of other domains. In addition, insertions of 418 bp and 5849 bp in intron 2, and of 1384 bp and 197 bp in introns 5 and 7 respectively, are found in *Ceacam2* compared with *Ceacam1* (Table 3 and Figure 1C).

Expression of Ceacam1 and Ceacam2 mRNAs in mouse tissues

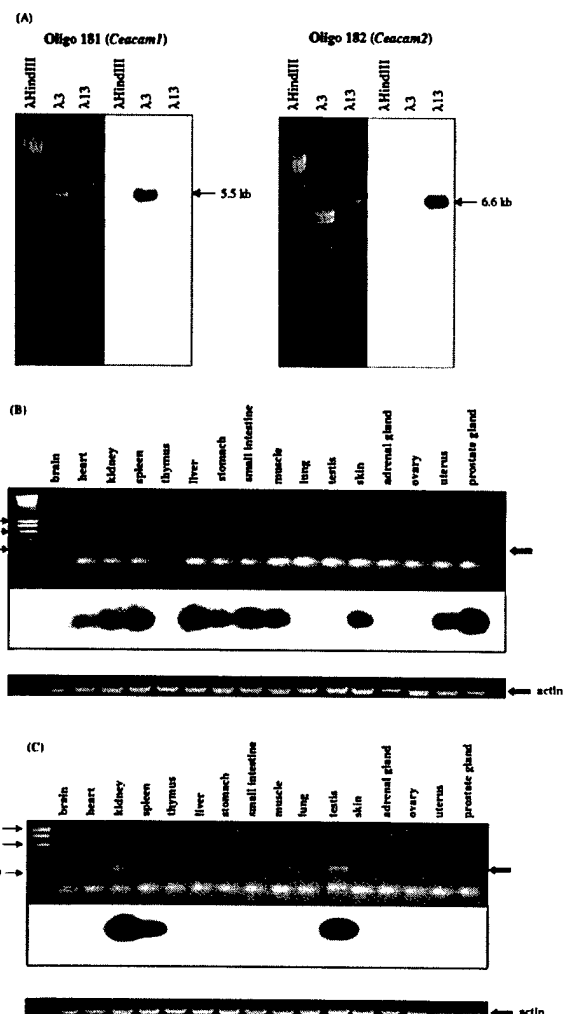To determine the tissue-specific distribution of *Ceacam1* and *Ceacam2*, we performed semi-quantitative RT-PCR on a panel of mouse cDNAs from various tissues by using *Ceacam1*- and *Ceacam2*-specific oligonucleotide pairs. The PCR products were hybridized further with *Ceacam1*- and *Ceacam2*-specific oligonucleotides derived from exon 2, the least similar exon (Figure 2). The specificities of these two oligonucleotides were confirmed by Southern blot analysis. As shown in Figure 3(A), Oligo 181 hybridized to DNA from λ3 (*Ceacam1*) but not from λ13 (*Ceacam2*), whereas Oligo 182 only hybridized to λ13 DNA. Both *Ceacam1* and *Ceacam2* mRNAs were detected, although not in the same tissues. This observation supports the notion that both *Ceacam1* and *Ceacam2* are expressed in mouse (Figures 3B and 3C). High levels of *Ceacam1* mRNA were detected in liver, small intestine, prostate and spleen (Figure 3B), similar to the tissue expression pattern in humans [22] and rats [23,24]. *Ceacam1* mRNA was also detected in heart, kidney, stomach, muscle, skin and uterus. In contrast, *Ceacam2* mRNAs were detected only in kidney, testis and, to a lesser extent, spleen (Figure 3C). These
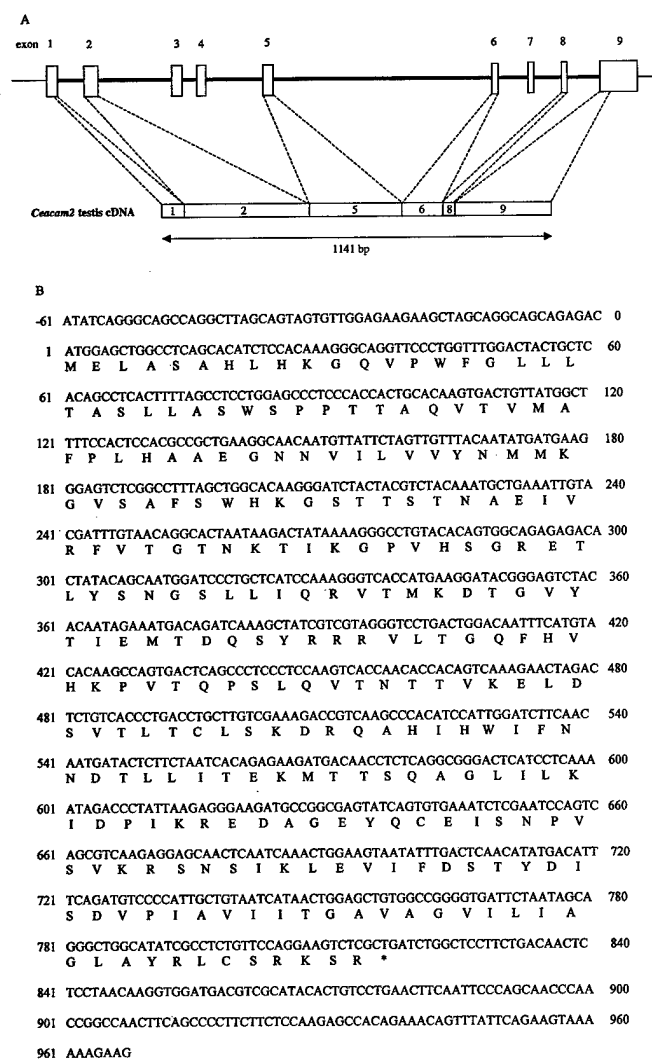
A

exon 1 2   3 4   5   6 7 8   9

Ceacam2 testis cDNA   1   2   5   6 8   9

1141 bp

B

-61 ATATCAGGGCAGCCAGGCTTAGCAGTAGTGTTGGAGAAGAAGCTAGCAGGCAGCAGAGAC 0

1 ATGGAGCTGGCCTCAGCACATCTCCACAAAGGGCAGGTTCCCTGGTTTGGACTACTGCTC 60
  M E L A S A H L H K G Q V P W F G L L L

61 ACAGCCTCACTTTTAGCCTCCTGGAGCCCTCCCACCACTGCACAAGTGACTGTTATGGCT 120
   T A S L L A S W S P P T T A Q V T V M A

121 TTTCCACTCCACGCCGCTGAAGGCAACAATGTTATTCTAGTTGTTTACAATATGATGAAG 180
    F P L H A A E G N N V I L V V Y N M M K

181 GGAGTCTCGGCCTTTAGCTGGCACAAGGGATCTACTACGTCTACAAATGCTGAAATTGTA 240
    G V S A F S W H K G S T T S T N A E I V

241 CGATTTGTAACAGGCACTAATAAGACTATAAAAGGGCCTGTACACAGTGGCAGAGAGACA 300
    R F V T G T N K T I K G P V H S G R E T

301 CTATACAGCAATGGATCCCTGCTCATCCAAAGGGTCACCATGAAGGATACGGGAGTCTAC 360
    L Y S N G S L L I Q R V T M K D T G V Y

361 ACAATAGAAATGACAGATCAAAGCTATCGTCGTAGGGTCCTGACTGGACAATTTCATGTA 420
    T I E M T D Q S Y R R R V L T G Q F H V

421 CACAAGCCAGTGACTCAGCCCTCCCTCCAAGTCACCAACACCACAGTCAAAGAACTAGAC 480
    H K P V T Q P S L Q V T N T T V K E L D

481 TCTGTCACCCTGACCTGCTTGTCGAAAGACCGTCAAGCCCACATCCATTGGATCTTCAAC 540
    S V T L T C L S K D R Q A H I H W I F N

541 AATGATACTCTTCTAATCACAGAGAAGATGACAACCTCTCAGGCGGGACTCATCCTCAAA 600
    N D T L L I T E K M T T S Q A G L I L K

601 ATAGACCCTATTAAGAGGGAAGATGCCGGCGAGTATCAGTGTGAAATCTCGAATCCAGTC 660
    I D P I K R E D A G E Y Q C E I S N P V

661 AGCGTCAAGAGGAGCAACTCAATCAAACTGGAAGTAATATTTGACTCAACATATGACATT 720
    S V K R S N S I K L E V I F D S T Y D I

721 TCAGATGTCCCCATTGCTGTAATCATAACTGGAGCTGTGGCCGGGGTGATTCTAATAGCA 780
    S D V P I A V I I T G A V A G V I L I A

781 GGGCTGGCATATCGCCTCTGTTCCAGGAAGTCTCGCTGATCTGGCTCCTTCTGACAACTC 840
    G L A Y R L C S R K S R *

841 TCCTAACAAGGTGGATGACGTCGCATACACTGTCCTGAACTTCAATTCCCAGCAACCCAA 900

901 CCGGCCAACTTCAGCCCCTTCTTCTCCAAGAGCCACAGAAACAGTTTATTCAGAAGTAAA 960

961 AAAGAAG

Figure 4 *Ceacam2*

(**A**) cDNA structure. The numbered boxes indicate exons. The thick lines connecting the exons represent introns. The broken lines show the portions of the exon sequences included in testis *Ceacam2* cDNA. (**B**) cDNA and protein sequences.

(A)   *Ceacam1*   (B)   *Ceacam2*

Figure 5 Expression of *Ceacam1* (A) and *Ceacam2* (B) mRNAs during embryonic development

(**A**) Analysis of RT-PCR products of *Ceacam1* on an agarose gel and gene-specific hybridization using Oligo 181. (**B**) Analysis of RT-PCR products of *Ceacam2* on an agarose gel and gene-specific hybridization using Oligo 182. Marker sizes are indicated by thin arrows, and the size of the PCR product is indicated by thick arrow.

and its sequence was determined (Figure 4). The mouse testis cDNA contained exons 1, 2, 5, 6, 8 and 9, and had an open reading frame of 273 amino acids (Figure 4B).

Expression of *Ceacam1* and *Ceacam2* during mouse embryonic development

The expression of *Ceacam1* and *Ceacam2* was examined in mouse embryos at days 8.5, 9.5, 12.5 and 19 of embryonic development by RT-PCR hybridization as described above. Using *Ceacam1*-specific Oligo 181, we observed age-related differences in mRNA levels: we detected the mRNA at day 8.5, it had disappeared at days 9.5–12.5, and it re-appeared at day 19 (Figure 5A). In contrast, no hybridization signal was detected when *Ceacam2*-specific Oligo 182 was used (Figure 5B). This result suggests that the expression of *Ceacam1*, but not that of *Ceacam2*, is developmentally regulated.

observations demonstrate that the *Ceacam1* and *Ceacam2* genes are differentially expressed in mouse tissues.

Ceacam2 cDNA

Although it was detected in testis, kidney and spleen, *Ceacam2* mRNA may not contain open reading frames for protein translation. As testis expressed only *Ceacam2*, the cDNAs coding for *Ceacam2* were obtained from testis RNA by RT-PCR with Oligo 630, which contained the 5′ untranslated region of exon 1, and Oligo 631, complementary to the exon 9 coding region at the termination codon TGA. The 1 kb PCR product obtained only hybridized to *Ceacam2*-specific Oligo 182, and not to *Ceacam1*-specific Oligo 181, suggesting that the fragment codes for *Ceacam2*. This result is consistent with the tissue distribution of *Ceacam2* (Figure 3C). The PCR product was then subcloned
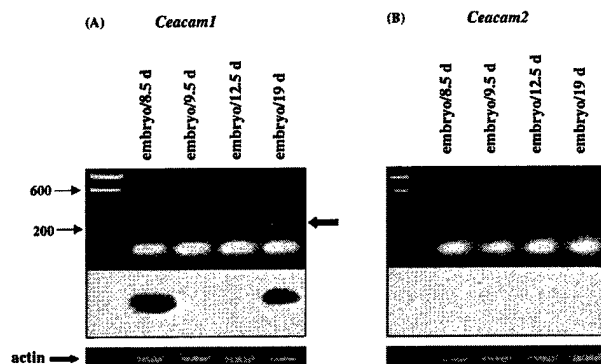
DISCUSSION

Because targeted gene deletion is performed in the 129 Sv mouse, we determined in the present studies the complete genomic structure and DNA sequences of the mouse *Ceacam1* and *Ceacam2* genes in a 129Sv/Ev mouse genomic library. We also examined the expression patterns of these two genes in adult mouse tissues and during embryonic development. Several conclusions can be made from this study. First, the *Ceacam1* and *Ceacam2* genes contain sufficient sequence differences that targeted gene deletion specific to either gene should be feasible. Secondly, functional redundancy may not be a problem when only one of these two genes is deleted, because, although *Ceacam1* and *Ceacam2* are highly homologous, they have different tissue expression patterns. Thirdly, *Ceacam1* is probably more important than *Ceacam2* in the mouse, because the tissue expression pattern of *Ceacam1* is similar to those of the single *Ceacam* genes in humans and rats. Conservation of the expression profile among these different species suggests that the function of

Ceacam1 may be essential. Fourthly, *Ceacam2* plays no role in embryonic development; it is not expressed in the mouse embryo.

The results of the present study provide important information for the design of gene targeting strategies for functional studies of *Ceacam* genes, by either deleting or introducing mutations into the *Ceacam* genes in the mouse germline. The differences in the two DNA sequences can be used to target a specific gene. As *Ceacam1* lacks the 418 and 5849 bp insertions in *Ceacam2* intron 2 and the 1384 bp insertion present in *Ceacam2* intron 5 (Figure 1C), gene-targeting vectors containing introns 2 or 5 of *Ceacam1* or *Ceacam2* could be used to achieve selective homologous recombination in the desired gene. If mutant mice in which only one *Ceacam* gene is deleted survive, then they can be crossed further to generate mice deficient in both *Ceacam* genes. However, as *Ceacam1* and *Ceacam2* are both on chromosome 7, it is possible that they are too close to allow construction of double-gene-knockout mice by crossing the single-gene-knockout mice.

The existence of the highly similar second *Ceacam* gene, *Ceacam2*, in mice raises the possibility that *Ceacam2* is a pseudogene. However, our results indicate that *Ceacam2* is probably not a pseudogene, since (1) *Ceacam2* contains a complete set of exons and introns typical of a *Ceacam* gene; (2) *Ceacam2* is transcribed in mouse, as shown by the presence of *Ceacam2* mRNA in several tissues; and (3) *Ceacam2* mRNA contains an open reading frame of 273 amino acids in testis, as has been shown in CMT-93 mouse rectal carcinoma cells [17]. These findings raise the interesting question of the function of *Ceacam2*. One of the functions of CEACAM1 is inhibition of tumour growth [3,6]. This suggests that this protein may play an important role in regulating cell growth and differentiation. Structural and functional analyses of rat CEACAM1 have revealed that the tumour-suppressive function requires a long cytoplasmic domain, generated by alternative splicing [4,6,24]. Because the cytoplasmic domains of CEACAM1 and CEACAM2 are identical, CEACAM2 may also have growth-suppressive activity. However, the role of CEACAM2 in testis, which is composed of rapidly dividing cells in spermatogenesis, is not clear. Thus the *in vivo* function of CEACAM2 may not be revealed until its gene is deleted in the mouse.

Knocking out a gene may have no phenotypic effect if related genes have similar functions. We showed that *Ceacam1* and *Ceacam2* are expressed in different tissues in the mouse. *Ceacam1* mRNA was detected in tissues rich in epithelial cells, which is consistent with expression of the CEACAM1 homologue in rats and humans [25–27]. In contrast, *Ceacam2* mRNA was abundantly expressed in testis, which does not express *Ceacam1* mRNA. In addition, *Ceacam2* mRNA was undetectable in the mouse embryo, whereas *Ceacam1* mRNA was developmentally regulated. Thus it appears that the two mouse *Ceacam* genes are not functionally redundant. However, we cannot rule out the possibility that the loss of *Ceacam1* expression may up-regulate expression of *Ceacam2*. These studies will have to await targeted gene deletion of these two related genes.

REFERENCES

1   Beauchemin, N., Draber, P., Dveksler, G., Gold, P., Gray-Owen, S., Grunert, F., Hammarstrom, S., Holmes, K. V., Karlsson, A., Kuroki, M. et al. (1999) Redefined nomenclature for members of the carcinoembryonic antigen family. Exp. Cell Res. **252**, 243–249

2   Lin, S. H., Culic, O., Flanagan, D. and Hixson, D. C. (1991) Immunochemical characterization of two isoforms of rat liver ecto-ATPase that show an immunological and structural identity with a glycoprotein cell-adhesion molecule with M_r 105000. Biochem. J. **278**, 155–161

3   Hsieh, J. T., Luo, W., Song, W., Wang, Y., Kleinerman, D. I., Van, N. T. and Lin, S. H. (1995) Tumor suppressive role of an androgen-regulated epithelial cell adhesion molecule (C-CAM) in prostate carcinoma cell revealed by sense and antisense approaches. Cancer Res. **55**, 190–197

4   Luo, W., Wood, C., Earley, K., Hung, M.-C. and Lin, S.-H. (1997) Suppression of tumorigenicity of breast cancer cells by an epithelial cell adhesion molecule (C-CAM1): The adhesion and growth suppression are mediated by different domains. Oncogene **14**, 1697–1704

5   Kleinerman, D. I., Dinney, C. P. N., Zhang, W.-W., Lin, S.-H., Van, N. T. and Hsieh, J.-T. (1996) Suppression of human bladder cancer growth by increased expression of C-CAM1 gene in an orthotopic model. Cancer Res. **56**, 3431–3435

6   Kunath, T., Ordonez-Garcia, C., Turbide, C. and Beauchemin, N. (1995) Inhibition of colonic tumor cell growth by biliary glycoprotein. Oncogene **11**, 2375–2382

7   Thompson, N. L., Lin, S. H., Panzica, M. A. and Hixson, D. C. (1994) Cell CAM 105 isoform RNA expression is differentially regulated during rat liver regeneration and carcinogenesis. Pathobiology **62**, 209–220

8   Neumaier, M., Paululat, S., Chan, A., Matthaes, P. and Wagener, C. (1993) Biliary glycoprotein, a potential human cell adhesion molecule, is down-regulated in colorectal carcinomas. Proc. Natl. Acad. Sci. U.S.A. **90**, 10744–10748

9   Nollau, P., Scheller, H., Kona-Horstmann, M., Rohde, S., Hagenmuller, F., Wagener, C. and Neumaier, M. (1997) Expression of CD66a (human C-CAM) and other members of the carcinoembryonic antigen gene family of adhesion molecules in human colorectal adenomas. Cancer Res. **57**, 2354–2357

10 Rosenberg, M., Nedellec, P., Jothy, S., Fleiszer, D., Turbide, C. and Beauchemin, N. (1993) The expression of mouse biliary glycoprotein, a carcinoembryonic antigen-related gene, is down-regulated in malignant mouse tissues. Cancer Res. **53**, 4938–4945

11 Bamberger, A. M., Riethdorf, L., Nollau, P., Naumann, M., Erdmann, I., Gotze, J., Brummer, J., Schulte, H. M., Wagener, C. and Loning, T. (1998) Dysregulated expression of CD66a (BGP, C-CAM), an adhesion molecule of the CEA family, in endometrial cancer. Am. J. Pathol. **152**, 1401–1406

12 Kleinerman, D. I., Troncoso, P., Lin, S. H., Pisters, L. L., Sherwood, E. R., Brooks, T., von Eschenbach, A. C. and Hsieh, J. T. (1995) Consistent expression of an epithelial cell adhesion molecule (C-CAM) during human prostate development and loss of expression in prostate cancer: implication as a tumor suppressor. Cancer Res. **55**, 1215–1220

13 Pu, Y. S., Luo, W., Lu, H.-H., Gingrich, J., Greenberg, N. and Lin, S.-H. (1999) Differential expression of C-CAM protein in prostate cancer progression in a transgenic adenocarcinoma mouse model. J. Urol. **162**, 892–896

14 Brandriff, B. F., Gordon, L. A., Tynan, K. T., Olsen, A. S., Mohrenweiser, H. W., Fertitta, A., Carrano, A. V. and Trask, B. J. (1992) Order and genomic distances among members of the carcinoembryonic antigen (CEA) gene family determined by fluorescence in situ hybridization. Genomics **12**, 773–779

15 Najjar, S. M., Boisclair, Y. R., Nabih, Z. T., Philippe, N., Imai, Y., Suzuki, Y., Suh, D. S. and Ooi, G. T. (1996) Cloning and characterization of a functional promoter of the rat pp120 gene, encoding a substrate of the insulin receptor tyrosine kinase. J. Biol. Chem. **271**, 8809–8817

16 Cheung, P. H., Culic, O., Qiu, Y., Earley, K., Thompson, N., Hixson, D. C. and Lin, S. H. (1993) The cytoplasmic domain of C-CAM is required for C-CAM-mediated adhesion function: studies of a C-CAM transcript containing an unspliced intron. Biochem. J. **295**, 427–435

17 Nedellec, P., Dveksler, G. S., Daniels, E., Turbide, C., Chow, B., Basile, A. A., Holmes, K. V. and Beauchemin, N. (1994) Bgp2, a new member of the carcinoembryonic antigen-related gene family, encodes an alternative receptor for mouse hepatitis viruses. J. Virol. **68**, 4525–4537

18 Williams, R. K., Jiang, G.-S. and Holmes, K. V. (1991) Receptor for mouse hepatitis virus is a member of the carcinoembryonic antigen family of glycoprotein. Proc. Natl. Acad. Sci. U.S.A. **88**, 5533–5536

19 Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd edn, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

20 Nedellec, P., Turbide, C. and Beauchemin, N. (1995) Characterization and transcriptional activity of the mouse biliary glycoprotein 1 gene, a carcinoembryonic antigen-related gene. Eur. J. Biochem. **231**, 104–114

21 Senapathy, P., Shapiro, M. B. and Harris, N. L. (1990) Splice junctions, branch point sites, and exons: sequence statistics, identification, and applications to genome project. Methods Enzymol. **183**, 252–278

22 Prall, F., Nollau, P., Neumaier, M., Haubeck, H. D., Drzeniek, Z., Helmchen, U., Loning, T. and Wagener, C. (1996) CD66a (BGP), an adhesion molecule of the carcinoembryonic antigen family, is expressed in epithelium, endothelium, and myeloid cells in a wide range of normal human tissues. J. Histochem. Cytochem. **44**, 35–41

23 Odin, P., Asplund, M., Busch, C. and Obrink, B. (1988) Immunohistochemical localization of cell CAM 105 in rat tissues: appearance in epithelia, platelets, and granulocytes. J. Histochem. Cytochem. **36**, 729–739

24 Estrera, V. T., Luo, W., Phan, D., Hixson, D. and Lin, S.-H. (1999) The cytoplasmic domain of C-CAM1 is necessary and sufficient in suppression of the growth of prostate cancer cells. Biochem. Biophys. Res. Commun. **263**, 797–803

25 Cheung, P. H., Thompson, N. L., Earley, K., Culic, O., Hixson, D. and Lin, S. H. (1993) Cell-CAM105 isoforms with different adhesion functions are coexpressed in adult rat tissues and during liver development. J. Biol. Chem. **268**, 6139–6146

26 Thompson, J. A., Grunert, F. and Zimmermann, W. (1991) Carcinoembryonic antigen gene family: molecular biology and clinical perspectives. J. Clin. Lab. Anal. **5**, 344–366

27 Hinoda, Y., Neumaier, M., Hefta, S. A., Drzeniek, Z., Wagener, C., Shively, L., Hefta, L. J. F., Shively, J. E. and Paxton, R. (1988) Molecular cloning of a cDNA coding biliary glycoprotein I: Primary structure of a glycoprotein immunologically crossreactive with carcinoembryonic antigen. Proc. Natl. Acad. Sci. U.S.A. **85**, 6959–6963